植 物 分 类 与 资 源 学 报 2014, **36** (6): 715~722
Plant Diversity and Resources DOI: 10.7677/ynzwyj201414036

蕨类植物问荆 DEAD-box RNA 解旋酶基因的克隆与分析*

徐道兰, 曹建国, 王全喜, 戴锡玲**
(上海师范大学生命与环境科学学院, 上海 200234)

摘要: DEAD-box RNA 解旋酶参与 RNA 多方面的代谢, 在植物生长发育和逆境反应中起重要作用。本研究从蕨类植物问荆 (*Equisetum arvense*) 中克隆到一条 DEAD-box RNA 解旋酶 cDNA 全长序列, 命名为 *EaRH1*, 并在 GenBank 注册登记 (KJ734026)。序列分析显示: 该 cDNA 全长 3 230 bp, 包含一个从 487 bp 到 2 799 bp 编码 770 个氨基酸的开放读码框, 其对应的蛋白序列包含 9 个保守模块结构。*EaRH1* 与其它物种 DEAD-box RNA 解旋酶蛋白序列比对结果显示: 模块 Ⅰa、Ⅱ和Ⅲ序列几乎完全相同, 模块 Q、Ⅰ和 Ⅳ序列存在一些差异。*EaRH1* 与江南卷柏 (*Selaginella moellendorffii*) 基因组一条假定序列相似度高达 69%, 其中相似度最高的区域集中在包含 9 个保守模块的结构域。系统进化树分析显示: *EaRH1* 与拟南芥 (*Arabidopsis thaliana*) DEAD-box RNA 解旋酶 *At3g22320* 在氨基酸序列上有相对较高的同源性。序列结构比较和进化分析可推测出 *EaRH1* 可能参与植物体生长发育、miRNA 生物合成、与 RNA 结合蛋白的相互作用和非生物胁迫应答。本文的研究为探索问荆 DEAD-box RNA 解旋酶的进一步功能提供参考。

关键词: 问荆; DEAD-box RNA 解旋酶; *EaRH1*
中图分类号: Q 785 **文献标识码**: A **文章编号**: 2095-0845(2014)06-715-08

Cloning and Characterization of DEAD-box RNA Helicases Gene from the Fern *Equisetum arvense*

XU Dao-Lan, CAO Jian-Guo, WANG Quan-Xi, DAI Xi-Ling**
(*College of Life and Environmental Science*, *Shanghai Normal University*, Shanghai 200234, China)

Abstract: DEAD-box RNA helicases play an important role in multiple biological processes, including growth, development, stress response and metabolism of cell. A full-length cDNA of DEAD-box RNA helicase, named *EaRH1*, has been cloned from *Equisetum arvense* in this study (GenBank accession number: KJ734026). Sequence analysis indicates that the cDNA of *EaRH1* is 3 230 bp in full length, and contains an ORF region from 487 bp to 2 799 bp, encoding 770 amino acids. The corresponding protein includes 9 conserved mode structures. Comparison with the sequence of the DEAD-box RNA helicase of other species, it is showed that the sequence of Motif Ⅰa, Motif Ⅱ and Motif Ⅲ are almost completely same while there are certain differences in sequence of Q motif, Motif Ⅰ and Motif Ⅳ. The similarity of the *EaRH1* sequence with the hypothetical DEAD-box RNA helicase protein of *Selaginella moellendorffii* reaches 69%. And the highest sequence similarity is concentrated in the domain that includes 9 conserved motifs. Phylogenetic analysis of DEAD-box RNA helicases sequences from different plants indicates that *Equisetum arvense* is more closely related to a DEAD-box gene of *Arabidopsis thaliana* (*At3g22320*). It can be inferred from sequence comparison and phylogenetic analysis that *EaRH1* might be involved in growth and development, miRNA biogenesis, interaction with RNA binding protein and abiotic stress responses. Our study provides references for further functional researches on DEAD-box RNA helicases of *Equisetum arvense*.

Key words: *Equisetum arvense*; DEAD-box RNA helicase; *EaRH1*

* Funding: 上海师范大学理科科研基金 (SK201231)
** Author for correspondence; E-mail: daixiling2010@shnu.edu.cn
Received date: 2014-03-07, Accepted date: 2014-05-27
作者简介: 徐道兰 (1987-) 女, 硕士研究生, 从事蕨类植物分子生物学研究。

Helicases are widely present in almost all life forms from virus to humans (Hail and Matin, 1999), which are divided into 5 superfamilies, SF1 - 5. DEAD-box RNA helicaces are members of SF2 family (Cordin *et al.*, 2006), which are subdivided into DEAD, DEAH, DECH and DEVH according to sequence differences of conserved DEAD motif. The DEAD-box RNA helicase family has been defined by Linder *et al.* (1989) and named according to the highly conserved residues, Asp-Glu-Ala-Asp, in motif Ⅱ. The DEAD-box family is characterized by the presence of nine conserved motifs which are involved in ATPase and helicase activities and in their regulation of many important life activities (Cordin *et al.*, 2004).

DEAD-box RNA helicases have profound effects on RNA metabolism, including RNA transcription, premRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation and RNA decay (Zhang *et al.*, 2006; Roack and Linder, 2004). Recent evidence indicates that DEAD-box RNA helicases are not only involved in growth and development of plants, translation initiation, cell proliferation, vegetative organs development, sexual reproduction, but also in plant abiotic stress responses, including cold stress, heat stress, salt stress, osmotic stress and oxidative stress (Yang *et al.*, 2009; Gong *et al.*, 2002, 2005).

So far, many DEAD-box RNA helicases in higher plants have been identified (Boudet *et al.*, 2001; Aubourg *et al.*, 1999; Mingam *et al.*, 2004; Lorkovic *et al.*, 1997; Owttrim *et al.*, 1991, 1994; Roel and Kuhlemeier, 1998; Gallie *et al.*, 1997; Webster *et al.*, 1991; Gendra *et al.*, 2004; Kato *et al.*, 2001; Nakamura *et al.*, 2004; Vashisht and Tuteja, 2005). Fifty-eight DEAD-box RNA helicases, designed as *AtRH1-AtRH58*, have been found in the model plant *Arabidopsis thaliana* which are almost evenly distributed in 5 chromosomes of *Arabidopsis thaliana* (Boudet *et al.*, 2001; Aubourg *et al.*, 1999; Mingam *et al.*, 2004). DEAD-box genes have been found also in *Spinacia oleracea* (Lorkovic *et al.*, 1997), *Nicotiana plumbaginifolia* (Owttrim *et al.*, 1991, 1994; Roel and Kuhlemeier, 1998; Wang *et al.*, 2000), *Triticum aestivum* (Gallie *et al.*, 1997), *Zea mays* (Zhang *et al.*, 2006; Webster *et al.*, 1991; Gendra *et al.*, 2004), *Oryza sativa* (Kato *et al.*, 2001; Li *et al.*, 2008), *Vigna radiate* (Li *et al.*, 2001), *Hordeum vulgare* (Nakamura *et al.*, 2004), *M. sativa* (Luo *et al.*, 2009) and *Pisum sativum* (Vashisht *et al.*, 2005). However, the DEAD-box genes in ferns are till poorly known. Here, we describe the cloning and characters of this novel DEAD-box RNA helicase in *Equisetum arvense*, aiming at revealing *Equisetum arvense* DEAD-box RNA helicases structures and provide references for the future functional studies.

1 Materials and methods

1.1 Materials

The reproductive stems of *Equisetum arvense* were collected in Tonghe county, Heilongjiang province in late April. The fresh materials were quick-freezed in liquid nitrogen and stored in −80 ℃ refrigerator.

1.2 RNA extraction and cDNA synthesize

The frozen materials were ground to fine powder in liquid nitrogen, and total RNA was extracted using the SV Total RNA Isolation Kit (Promega) according to the manufacturer's protocol. Total RNA was treated with Recombinant DNase Ⅰ (Takara) for 30 min at 37 ℃. The integrity and purity of the samples were assessed by agarose gel electrophoresis and spectrophotometer quantification. All of them presented high degree of integrity and purity ($1.8 < A_{260}/A_{280} < 2.0$); thus, they were used later for the synthesis of cDNA single strand. cDNAs were synthesized by using the Superscript Ⅲ reverse transcriptase (Invitrogen) and the adaper primer (5′-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT-3′) supplied by Invitrogen.

1.3 Isolation of DEAD-box fragments by PCR

Degenerate primers were designed based on the reported sequences in Genbank. The following forward primer was used: F1 (5′-CCWCNDCSCWNT-

KCASDGWKCWK-3′). The following reverse primer was used: R1 (5′-CAWTSTWSTYYATNSMTSNDYAT-3′). PCR was carried out using 1 μL cDNA and 0.2 μm specific primers in a 50 μL reaction volume containing standard buffer (Mg^{2+} plus), 2.5 mmol·L^{-1} of each dNTPs and 2.5 U Taq polymerase and performed at 94 ℃ for 5 min, followed by 35 cycles of 45 sec at 94 ℃, 45 sec at 55 ℃, and 1 min at 72 ℃, and a final extension of 10 min at 72 ℃. PCR products were separated on 2% agarose gels for electrophoresis and stained with ethidium bromide. The single specific band of PCR products was cloned into the pEASY-T5 zero cloning kit (Trans) for sequencing.

1.4 Isolation of DEAD-box containing cDNAs using RACE

Total RNA was extracted from the strobile of *Equisetum arvense*. 3′-RACE and 5′-RACE were both performed by SMARTer RACE cDNA Amplification Kit (Clontech) according to the manufacturer's protocol. The primers used in 3′-RACE were F2 (5′-TGCGGATGATGTGGCCTCTAC-3′) and F3 (5′-CGGTTTCCGTGAGGGCAGATT-3′). The primers used in 5′-RACE were R2 (5′-CCGTTGGGATGGAGAAGAGTTT-3′) and R3 (5′-ACATCTTCAGCGAAACCAACCC-3′). The PCR condition was 94 ℃ 2 min, followed by 25 cycles of 94 ℃ for 30 second, 68 ℃ for 30 second, and 72 ℃ for 3 min, and a final step at 72 ℃ for 6 min using Advantage 2 Polymerase Mix (Takara). PCR products were separated on 1.2% agarose gels and the single specific band of PCR products was cloned into the pEASY-Blunt zero cloning kit (Trans) for sequencing. Because of potential PCR errors, at least two independently amplified clones were sequenced in every experiment. ORF amplification was processed by specific primers designed according to 3′- and 5′-RACE results and performed according to the KOD polymerase provided by TOYOBO.

1.5 Sequencing and phylogenetic analysis

Multiple sequence alignments were performed on the DEAD-box protein sequences using Clustal W with default parameters, and the alignments were then adjusted manually. A phylogenetic tree was constructed with aligned DEAD-box protein sequences using MEGA 5.2 by employing the neighbor joining (NJ) method with the following parameters: Jones-Taylor-Thornton (JTT) model, complete deletion, and bootstrap (1 000 replicates).

2 Results

2.1 Cloning of *EaRH1*, an *Equisetum arvense* gene encoding a DEAD-box RNA helicase

This study clones an amplification product about 1 285 bp using degenerate primers, F1 and R1 (Fig. 1A). Based on this fragment, RACE method is performed and works out its complete 5′ and 3′ terminals (Fig. 1B, C). Results show that this cDNA sequence has poly A structure, which indicates that this clone contains complete 3′-terminal domain. This cDNA, named *EaRH1* (GenBank accession number: KJ734026), 3 230 bp in full length, contains an ORF region from 487 bp to 2 799 bp, encoding 770 amino acids (Fig. 1D; Fig. 2).

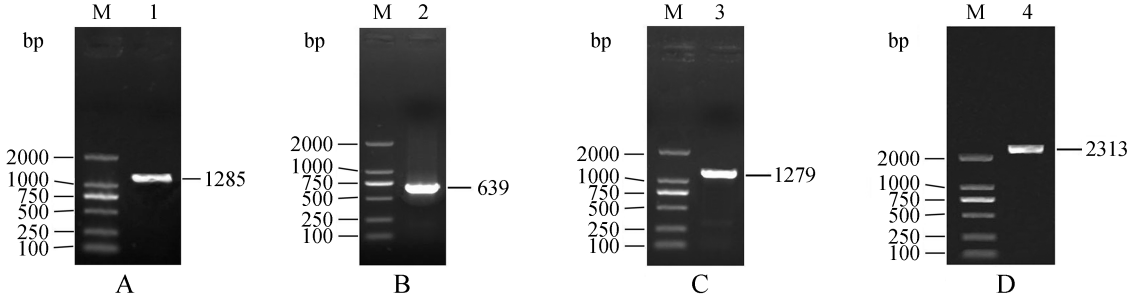

Fig. 1 Electrophoresis results of PCR amplification of *Equisetum arvense* DEAD-box RNA helicase *EaRH1* cDNA ends

M. DL2000 DNA marker; A. The result of core fragment PCR (lane 1); B. The result of 5′ RACE PCR (lane 2); C. The result of 3′ RACE PCR (lane 3); D. Amplification of the coding region of *EaRH1* (lane 4)

```
1    ctaaaaatatgtttggccttttttaaacaaagcgcatgaaacagtttgcactttgagtttg
61   gagggtttttattgcatccttctctctctctcctctctytctctctcattctctctctct
121  ctctctctggttcctccggcaggtcatggtcatggtcatgacctaccgaaggttttccta
181  aacaagttctctttttctccttctctttttttctctctttcctctttttctctcyctctct
241  ctaactctctctgtctcggaagtccatttgggaagacctcgctttggggttttcgcctg
301  attcgatcgacccgagcccctcgtcgtccgcgggttggccttgcgacttcgggtttccctt
361  ccgttggcgccggcggtagtcgaaggcattgccaatgggagatgctgccccccccattcc
421  cccactggaaacaacaagccccatggagttgggggtttcgctggtctccttcctctcggcc
481  cgccccatgtccgttggcgtggaggcacttcgggaggatgatgctgtcccggaagaaggg
                M  S  V  G  V  E  A  L  R  E  D  D  A  V  P  E  E  G
541  agtcctcgttccgaatcaatgtcgcagggccttgaaatatctaagattggagttgaacaa
       S  P  R  S  E  S  M  S  Q  G  L  E  I  S  K  I  G  V  E  Q
601  gagattgttgacgcccttgctcgtcgttctattgtccgcctctttccctattcagggtgct
       E  I  V  D  A  L  A  R  R  S  I  V  R  L  F  P  I  Q  G  A
661  gtactcgagccagcaatgaaaggacaagatttgattgggcgagcaaagacaggcacagga
       V  L  E  P  A  M  K  G  Q  D  L  I  G  R  A  K  T  G  T  G
721  aagacactcgcattcggaatcccccatcataaaccacgtcatcaaggaaaacaagggttca
       K  T  L  A  F  G  I  P  I  I  N  H  V  I  K  E  N  K  G  S
781  agaggaatgagaccatatggaaaacgaccgctagcactagttctagctcctacaagagag
       R  G  M  R  P  Y  G  K  R  P  L  A  L  V  L  A  P  T  R  E
841  cttgcgaagcaggtagaaaaagagttcaaagaatcggcgccaagtttgatgacggtttgc
       L  A  K  Q  V  E  K  E  F  K  E  S  A  P  S  L  M  T  V  C
901  gtgtatggagggacttcaattattatgcagcagcaacaattgcaaagaggggtggacata
       V  Y  G  G  T  S  I  I  M  Q  Q  Q  Q  L  Q  R  G  V  D  I
961  gctgttggtactccgggtcgtgtcattgacctcctcgaaagaggctccttgaatttgggt
       A  V  G  T  P  G  R  V  I  D  L  L  E  R  G  S  L  N  L  G
1021 gaagtgaaatttgttgtgctggatgaagctgatcagatgttgcgggttggtttcgctgaa
       E  V  K  F  V  V  L  D  E  A  D  Q  M  L  R  V  G  F  A  E
1081 gatgttgagaagatagcgggatatttgccagctgtgcggcaaagcatgctcttttcagca
       D  V  E  K  I  A  G  Y  L  P  A  V  R  Q  S  M  L  F  S  A
1141 acaatgccggactgggtcaagaatatctcaaggaaatatttgaagaatcctactattgtc
       T  M  P  D  W  V  K  N  I  S  R  K  Y  L  K  N  P  T  I  V
1201 gatctggtcggtgagaacgatgaaaaattagcagaggggataaaactctttctccatccca
       D  L  V  G  E  N  D  E  K  L  A  E  G  I  K  L  F  S  I  P
1261 acggtgaattcagcaaaacgtttctgtttctaagcgacctaatagctgttttatgggaaaagt
       T  V  N  S  A  K  R  S  V  L  S  D  L  I  A  V  Y  G  K  S
1321 aaaaaaaccattgttttcactcaaacaaaacgcgatgcggatgatgtggcctctactttg
       K  K  T  I  V  F  T  Q  T  K  R  D  A  D  D  V  A  S  T  L
1381 gggcgagctcatgcatgtgaggctttacatggtgacatcacgcaagcccaacgagagaag
       G  R  A  H  A  C  E  A  L  H  G  D  I  T  Q  A  Q  R  E  K
1441 acactgaacggtttccgtgagggcagattcaaggttctggttgctactgatgttgcttcc
       T  L  N  G  F  R  E  G  R  F  K  V  L  V  A  T  D  V  A  S
1501 cgtgggcttgatgttcctaatgttgatctggttattcattatgagattcccaatgactca
       R  G  L  D  V  P  N  V  D  L  V  I  H  Y  E  I  P  N  D  S
1561 gaaacttttgtgcatcgatcgggaaggacagggcgagctggcaaggagggagttgccatt
       E  T  F  V  H  R  S  G  R  T  G  R  A  G  K  E  G  V  A  I
1621 ctgatgtatacactacagcaaaatcgtactatttcaatcattgagcgggatgttggatgc
       L  M  Y  T  L  Q  Q  N  R  T  I  S  I  I  E  R  D  V  G  C
1681 aaatttgagaaaattagccctcctagtgcagaggaagttatcaaatctacagctgatcag
       K  F  E  K  I  S  P  P  S  A  E  E  V  I  K  S  T  A  D  Q
1741 gcagtaggagcactccagaaggtccacccagacttgaggggtatatttttaccgattgcg
       A  V  G  A  L  Q  K  V  H  P  D  L  R  G  I  F  L  P  I  A
1801 gaaaaacttctaagtgagcaaggagcaggagcttttgcctccgcccttgcatgcttaagt
       E  K  L  L  S  E  Q  G  A  G  A  F  A  S  A  L  A  C  L  S
1861 gggtttacacagcctcctgcttcacgctcttctacttactcatgaacagggatggacaaca
       G  F  T  Q  P  P  A  S  R  S  L  L  T  H  E  Q  G  W  T  T
1921 ttgaggctttctcggccgaaaagcgggccgccattgttcaccactcaatcggttattgct
       L  R  L  S  R  P  K  S  G  P  P  L  F  T  T  Q  S  V  I  A
1981 gcattagctcaaatttgtccttctgcgactaatagagtgggaaaaattcaaatgctcaat
       A  L  A  Q  I  C  P  S  A  T  N  R  V  G  K  I  Q  M  L  N
2041 gaagtcgatgtccaaggagcggtatttgacatgcccgaaagcattgcaaaggagcttctt
       E  V  D  V  Q  G  A  V  F  D  M  P  E  S  I  A  K  E  L  L
2101 tccctgcagcataactcgagcgatacacttgatgctcccaagcagcttcctaatattgtt
       S  L  Q  H  N  S  S  D  T  L  D  A  P  K  Q  L  P  N  I  V
2161 gaagatatcccatctacgagttactatggtagatcgtttggaggcggcggaggtttggc
       E  D  I  P  S  T  S  Y  Y  G  R  S  F  G  G  G  G  G  F  G
2221 aacagacctagggtaggtccacctttaatggcaggaatgattcatatggcaggccatct
       N  R  P  R  V  G  P  P  F  N  G  R  N  D  S  Y  G  R  P  S
2281 acaggtgacagaagttctggtgaacgaccgtcattttatccgccaaataatggcacgagt
       T  G  D  R  S  S  G  E  R  P  S  F  Y  P  P  N  N  G  T  S
2341 gggccccgcatgtctacgccctatccaggggggccaggttctagtggggtacgtatgccc
       G  P  R  M  S  T  P  Y  P  G  G  P  G  S  S  G  V  R  M  P
2401 cccccgccatatttctggtggaaacggagtaagtgggttccgcatgtcggcgccatatccg
       P  P  P  Y  S  G  G  N  G  V  S  G  F  R  M  S  A  P  Y  P
2461 ccatatcctggaagagtagcggataatagctctagtgggttgcggccacccacaccctac
       P  Y  P  G  R  V  A  D  N  S  S  S  G  L  R  P  P  T  P  Y
2521 tcatatgggtccaacacgggaaccgggggaccgagtaataggaggatcagcaggctctttt
       S  Y  G  S  N  T  G  T  G  D  R  V  I  G  G  S  A  G  S  F
2581 gtaagttcgcttgggagaaacggagcatccttcaatgaagacagaggggtggcttagca
       V  S  S  L  G  R  N  G  A  S  F  N  E  D  R  G  G  G  L  A
2641 gactttggraggtctctttctgagaggaaatttgatagcggcgttgggggcattggttcc
       D  F  G  R  S  L  S  E  R  K  F  D  S  G  V  G  G  I  G  S
2701 tcggattccaccctcttgaatgacaagtggaacaaatgtcrgctatgcggacaacccggt
       S  D  S  T  L  L  N  D  K  W  N  K  C  X  L  C  G  Q  P  G
2761 catttcgctcgcaattgcccaactctgcagaaaaagttagagagggtggcaaacttgctgg
       H  F  A  R  N  C  P  T  L  Q  K  S  *
2841 ratcgtgaatatccccttcacatgtggagagcttcatgcacttcacaggtatgcctgcca
2901 ttagtaagggatttttaaaatgcactctgataatgtagaggtagtatgctaaaccaaggg
2961 ggargggggagaagggkgaagtayataggctcctgtcttaatgtgtaaaagctgaygcc
3021 ttctaagatgccygggtataccgagattatttcagtcatgtgatggaatgtgatataaagg
3061 gagcctagaagtgtagactctttcgattttatctgtaaagcatcgcactttgtaatgtgc
3121 tttaaatactaagattaytgtgggacactaggtgacgattaggtttcaatacagctatcgt
3181 gatgatgcctgctttagtgagtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Fig. 2 Nucleotides sequence and deduced amino acids sequence of *Equisetum arvense* DEAD-box RNA helicase *EaRH1*.

The translational start coden (ATG) is in bold and the stop coden is indicated by an asterisk

2.2 The structure of *EaRH1* protein and phylogenetic analysis

Searching for conserved domains revealed that *EaRH1* protein contains two typical conserved domains, DEADc and HELICc (helicase superfamily C-terminal domain), which are characteristic of the DEAD-box protein family (Fig. 3). Besides, *EaRH1* also contains 9 conserved mode structures of reported DEAD-box RNA helicases by amino acids sequence BLAST comparison (Fig. 4A).

DEAD-box RNA helicases contain many conserved motif structures, which are associated with biochemical functions. Comparision of *EaRH1* with DEAD-box RNA helicases conserved motifs so far identified reveals that their amino acids sequences of conserved domain are highly identical (Fig. 4B). This also indicates that *EaRH1* belongs to the family and have biochemical functions similar to DEAD-box RNA helicases from other species. However, there are still certain differences between them, Q motif, Motif Ⅰ and Motif Ⅳ, for example, so *EaRH1* protein is a new DEAD-box RNA helicase.

In ferns, only *Selaginella moellendorffii* has completed the whole genome sequencing by far. Sequence comparison result shows that the similarity *EaRH1* and a hypothetical DEAD-box RNA helicase protein from *Selaginella moellendorffii* reaches 69% (Fig. 5). The highest sequence similarity is concentrated in the domain that includes 9 conserved motifs.

To research the evolutionary relationship, phylogenetic tree is constructed on *EaRH1* and DEAD-box RNA helicase proteins from other species (Fig. 6). Result shows that *EaRH1* forms a branch with *Arabidopsis thaliana At3g22320*. Genetic distance and phylogenetic relationship indicate that *EaRH1* is a new DEAD-box RNA helicase. Besides, *EaRH1* also shares a relatively high similarity with *Ricinus communis Rc-XM002512556* and *Hordeum vulgare HVD1*. This also indicates that DEAD-box RNA helicase motif structures are relatively conserved during plant evolution.

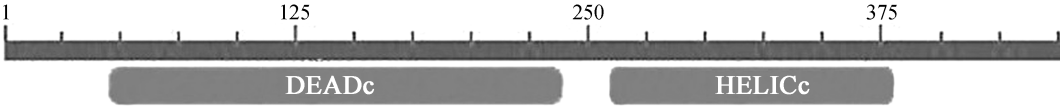

Fig. 3　The NCBI conserved domain search result of the deduced amino acids sequence of *Equisetum arvense* DEAD-box RNA helicase *EaRH1*
DEADc: DEAD-box helicases; HELICc: Helicase superfamily C-terminal domain

Fig. 4　Structural organization of the predicted *EaRH1* protein
A. DEAD-box RNA helicates' conserved motifs in deduced amino acids sequence of *Equisetum arvense* DEAD-box RNA helicase *EaRH1*. The nine conserved motifs of DEAD-box RNA helicases are indicated by underline. B. Comparison of *EaRH1* and other conserved motifs of DEAD-box RNA helicases

```
EaRH1                      26   SQGLEISKIGVEQEIVDALARRSIVRLFPIQGAVLEPAMKGQDLIGRAKTGTGKTLAFGI   85
                                S+ L I+ +G++  IVDALA+R IV LFPIQ AV +PAM+G DLI RAKTGTGKTLAFGI
Selaginella moellendorffii      1   SEELAIASLGIDSRIVDALAKRGIVNLFPIQRAVFQPAMQGLDLIARAKTGTGKTLAFGI   60

EaRH1                      86   PIINHVIKENKGSRGMRPYGKRPLALVLAPTRELAKQVEKEFKESAPSLMTVCVYGGTSI   145
                                PI++++ +E K S   RP G  P A++LAPTRELAKQVE E   +AP L  VC+YGG SI
Selaginella moellendorffii      61   PILDNISRE-KSSMQFRPAGT-PRAIILAPTRELAKQVENELVLAAPHLSVVCIYGGVSI   118

EaRH1                      146  IMQQQQLQRGVDIAVGTPGRVIDLLERGSLNLGEVKFVVLDEADQMLRVGFAEDVEKIAG   205
                                  Q++QL   V++ VGTPGRVID+LERG L L EVKF +LDEADQML VGFAE+VE+I
Selaginella moellendorffii      119  EGQRRQLASSVEVVVGTPGRVIDMLERGDLRLNEVKFAILDEADQMLAVGFAEEVERIMQ   178

EaRH1                      206  YLPAVRQSMLFSATMPDWVKNISRKYLKNPTIVDLVGENDEKLAEGIKLFSIPTVNSAKR   265
                                  LP  R +MLFSATMP W++N++ KYLKNPT++DLVGEN++K+AEGIKL+++ T + AK
Selaginella moellendorffii      179  KLPLQRHTMLFSATMPSWIRNLTSKYLKNPTMIDLVGENEDKVAEGIKLYAVSTNDLAKN   238

EaRH1                      266  SVLSDLIAVYGKSKKTIVFTQTKRDADDVASTLGRAHACEALHGDITQAQREKTLNGFRE   325
                                +L+DL+ VY K  K IVFT+TKRD DDVA+ + R+  CEALHGDI+Q QREKTL+GFR+
Selaginella moellendorffii      239  KLLADLLTVYAKGGKAIVFTKTKRDTDDVAAVMSRSIGCEALHGDISQYQREKTLSGFRD   298

EaRH1                      326  GRFKVLVATDVASRGLDVPNVDLVIHYEIPNDSETFVHRSGRTGRAGKEGVAILMYTLQQ   385
                                GRF VLVATDVA+RGLD+PNVDL++HYE+P DSETFVHRSGRTGRAGK+G  ILMYTL Q
Selaginella moellendorffii      299  GRFNVLVATDVAARGLDIPNVDLIVHYEVPGDSETFVHRSGRTGRAGKKGTCILMYTLNQ   358

EaRH1                      386  NRTISIIERDVGCKFEKISPPSAEEVIKSTADQAVGALQKVHPDLRGIFLPIAEKLLSEQ   445
                                R +  IE +VGCKFE I PP+AE+V+ S++D A  +++ VHPDL  +FLP A+KLL EQ
Selaginella moellendorffii      359  KRVLRTIESEVGCKFEAIGPPAAEDVLASSSDHASESIKSVHPDLAKLFLPAAQKLLEEQ   418

EaRH1                      446  GAGAFASALACLSGFTQPPASRSLLTHEQGWTTLRLSRPKSGPPLFTTQSVIAALAQICP   505
                                G  A A+A+A +SG TQ P  RSLLT+E+G+ TL+L    +G  L  + +V +A+ Q+
Selaginella moellendorffii      419  GVNALAAAMAHVSGVTQLPPKRSLLTYEEGYMTLQL----TGGNLIDSSAVSSAVMQLSS   474

EaRH1                      506  SATNRVGKIQMLNEVDVQGAVFDMPESIAKELLSLQHNSSDTLDAPKQLPNIVED   560
                                 A   +G++ ML+     G VFD+PE +AKE+L+LQ+  S+ L A K LP++ ED
Selaginella moellendorffii      475  KAGMSMGRVTMLSSSTATGGVFDLPEDLAKEILALQNVGSNVLTAVKVLPDLQED   529
```

Fig. 5 Sequence comparison of *Equisetum arvense EaRH1* and a hypothetical DEAD-box RNA helicase protein from *Selaginella moellendorffii*

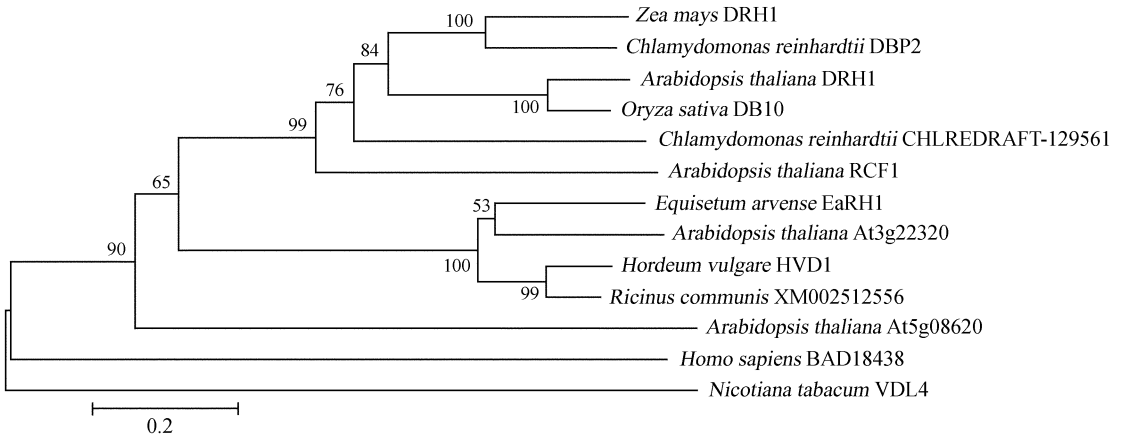

Fig. 6 Phylogenetic tree analysis of *Equisetum arvense* DEAD-box RNA helicase *EaRH1* with DEAD-box RNA helicase protein from other organisms

The DEAD-box RNA helicase proteins used were: *Equisetum arvense EaRH1*; *Arabidopsis thaliana At3g22320*, *RCF1*, *DRH1* and *At5g08620*; *Oryza sativa DB10*; *Homo sapiens BAD18438*; *Zea mays DRH1*; *Hordeum vulgare HVD1*; *Ricinus communis XM002512556*; *Nicotiana tabacum VDL4*; *Chlamydomonas reinhardtii CHLREDRAFT-129561* and *DBP2*

3 Discussion

The DEAD-box RNA helicase family was defined by Linder *et al.* (Linder *et al.*, 1989), who firstly found several highly conserved motifs existing in eight eukaryotic initiation factor *eIF4A* homologous proteins, mice *eIF-4AI*, *eIF-4AII*, *PL10*; yeast *TIF-4A*, *MSS116*; human *p68*; *Drosophila melanogaster vasa*; and *E. coli SrmB* (Yang *et al.*, 2009). All of these proteins contain 9 highly conserved motifs: Q motif, Motif Ⅰ, Motif Ⅰa, Motif Ⅰb, Motif Ⅱ, Motif Ⅲ, Motif Ⅳ, Motif Ⅴ, Motif Ⅵ (Linder *et al.*, 1989; Tanner *et al.*, 2003; Linder, 2006). The existence of all these nine motifs is the standard to determine whether a protein belongs to this family

(Linder *et al.*, 1989). *EaRH1* contains these conserved motif structures, so *EaRH1* of *Equisetum arvense* is a DEAD-box RNA helicase. Sequence comparison of *EaRH1* and a *Selaginella moellendorffii* hypothetical DEAD-box RNA helicase protein also shows that 9 conserved motifs are relatively conservative with plant evolution.

Many DEAD-box RNA helicases in higher plants have been identified by far. Besides, several DEAD-box RNA helicases of the algae *Chlamydomonas reinhardtii* are found in the NCBI database. However, no DEAD-box RNA helicase has been investigated in ferns. This study firstly clones *EaRH1*, one of *Equisetum arvense* DEAD-box RNA helicases, full-length cDNA sequence and enriching the DEAD-box RNA helicases family in plants. According to the phylogenetic tree analysis, *EaRH1* shares a relatively high similarity with *Arabidopsis thaliana At3g22320*, *Ricinus communis Rc-XM002512556* and *Hordeum vulgare HVD1*. However, due to the DEAD-box RNA helicases studies were relatively few, the phylogenetic tree is also not well reflecting the evolutionary relationships of the gene.

Comparison with the sequence of the DEAD-box RNA helicase of other species, it is showed that the sequence of Motif Ⅰa, Motif Ⅱ and Motif Ⅲ are almost completely same while there are certain differences in sequence of Q motif, Motif Ⅰ and Motif Ⅳ. The conserved motifs of DEAD-box RNA helicases proteins are associated with biochemical functions. Motif Ⅱ and motif Ⅵ are involved in the ATP hydrolysis, motif Ⅲ is related to RNA desmolysis (Cordin *et al.*, 2006). Motif Ⅱ and motif Ⅲ of *EaRH1* are DEAD and SAT respectively, which is consistent with DEAD group of DEAD-box RNA helicases proteins. Previous studies reveal that DEAD-box RNA helicases lose RNA desmolysis activity when mutation occurred in motif Ⅲ, that is SAT (Luking *et al.*, 1998). This conservatism might also imply that *EaRH1* has important biological functions, like the DEAD-box RNA helicases from other species.

The available evidences indicate that DEAD-box RNA helicases play an extensive and profound role in the life of plants (Yang *et al.*, 2009). Among the DEAD-box RNA helicases family, *eIF4A* is the typical representative of this family. Previous studies have confirmed that *Nicotiana tabacum* contains a number of *eIF4A* function as translation initiation (Owttrim *et al.*, 1991, 1994). Besides, in terms of plant growth and development, *PRH75* from *Spinacia oleracea* is related to the rapid cell growth and division (Lorkovic *et al.*, 1997), *CAF* (*DCL1*) from *Arabidopsis thaliana* is related to the floral meristem determinacy and miRNA biogenesis (Jacobsen *et al.*, 1999; Park *et al.*, 2002), *VrRH1* from *Vigna radiate* is related to the viability of mung bean seeds (Li *et al.*, 2001), and *ZmDRH1* from *Zea mays* is related to the interaction with RNA binding protein MA16 (Gendra *et al.*, 2004). In addition to these, DEAD-box RNA helicases are involved in plant abiotic stress responses. *FL2-5A4*, *LOS4* and *PMH2* from *Arabidopsis thaliana*, for example, are associated with cold responses (Gong *et al.*, 2002, 2005; Seki *et al.*, 2001; Matthes *et al.*, 2007). Based on the functional analysis of the DEAD-box RNA helicases in *Arabidopsis thaliana*, *Ricinus communis* and *Hordeum vulgare*, we also predict that *EaRH1* might be involved in growth and development, miRNA biogenesis, interaction with RNA binding protein and abiotic stress responses. Of course, further researches are needed to explore the function of the gene.

References:

Aubourg S, Kreis M, Lecharny A, 1999. The DEAD-box RNA helicase family in *Arabidopis thaliana* [J]. *Nucleic Acids Research*, **27**: 628—636

Boudet N, Aubourg S, Toffano-Nioche C *et al.*, 2001. Evolution of intron/exon structure of DEAD helicase family genes in *Arabidopsis*, *Caenorhabditis*, and *Drosophila* [J]. *Genome Research*, **11**: 2101—2114

Cordin O, Tanner NK, Doere M *et al.*, 2004. The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity [J]. *The EMBO Journal*, **23** (13): 2478—2487

Cordin O, Banroques J, Tanner NK *et al.*, 2006. The DEAD-box protein family of RNA helicases [J]. *Gene*, **367**: 17—37

Gallie DR, Le H, Caldwell C *et al.*, 1997. The phosphorylation state of translation initiation factors is regulated developmentally and following heat shock in wheat [J]. *Journal of Biological Chemistry*, **272** (2): 1046—1053

Gendra E, Moreno A, Alba MM *et al.*, 2004. Interaction of the plant glycine-rich RNA-binding protein *MA16* with a novel nucleolar DEAD box RNA helicase protein from *Zea mays* [J]. *The Plant Journal*, **38** (6): 875—886

Gong Z, Lee H, Xiong L *et al.*, 2002. RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance [J]. *Proceedings of the National Academy of Sciences of the United States of America*, **99** (17): 11507—11512

Gong Z, Dong CH, Lee H *et al.*, 2005. A DEAD-box RNA helicase is essential for mRNA export and important for development and stress responses in *Arabidopsis* [J]. *The Plant Cell*, **17**: 256—267

Hail MC, Matin SW, 1999. Heliease motifs: the engine that powers DNA unwinding [J]. *Molecular Microbiology*, **24** (5): 867—877

Jacobsen SE, Running MP, Meyerowitz EM, 1999. Disruption of an RNA helicase/RNAse Ⅲ gene in *Arabidopsis* causes unregulated cell division in floral meristems [J]. *Development*, **126** (23): 5231—5243

Kato A, Fujita S, Komeda Y, 2001. Identification and characterization of two genes encoding the eukaryotic initiation factor 4A in rice [J]. *DNA Sequence*, **12** (5-6): 295—303

Li DY, Liu HZ, Zhang HJ *et al.*, 2008. *OsBIRH1*, a DEAD-box RNA helicase with functions in modulating defence responses against pathogen infection and oxidative stress [J]. *Journal of Experimental Botany*, **59** (8): 2133—2146

Li SC, Chung MC, Chen CS, 2001. Cloning and characterization of a DEAD box RNA helicase from the viable seedlings of aged mung bean [J]. *Plant Molecular Biology*, **47** (6): 761—770

Linder P, Lasko PF, Ashburner M *et al.*, 1989. Birth of the D-E-A-D [J]. *Nature*, **337**: 121—122

Linder P, 2006. DEAD-box proteins: a family affair-active and passive players in RNP-remodeling [J]. *Nucleic Acids Research*, **34** (15): 4168—4180

Lorkovic ZJ, Herrmann RG, Oelmuller R, 1997. *PRH75*, a new nucleus-localized member of the DEAD-box protein family from higher plants [J]. *Molecular and Cellular Biology*, **17** (4): 2257—2265

Luking A, Stahl U, Schmidt U, 1998. The protein family of RNA helicases [J]. *Critical Reviews in Biochemistry and Molecular Biology*, **33**: 259—296

Luo Y, Liu YB, Dong YX *et al.*, 2009. Expression of a putative alfalfa helicase increases tolerance to abiotic stress in *Arabidopsis* by enhancing the capacities for ROS scavenging and osmotic adjustment [J]. *Journal of Plant Physiology*, **166**: 385—394

Matthes A, Schmidt-Gattung S, Kohler D *et al.*, 2007. Two DEAD-box proteins may be part of RNA-dependent high-molecular-mass protein complexes in *Arabidopsis* mitochondria [J]. *Plant Physiology*, **145**: 1637—1646

Mingam A, Toffano-Nioche C, Brunaud V *et al.*, 2004. DEAD-box RNA helicases in *Arabidopsis thaliana*: establishing a link between quantitative expression, gene structure and evolution of a family of genes [J]. *Plant Biotechnology Journal*, **2**: 401—415

Nakamura T, Muramoto Y, Yokota S *et al.*, 2004. Structural and transcriptional characterization of a salt-responsive gene encoding putative ATP-dependent RNA helicase in barley [J]. *Plant Science*, **167** (1): 63—70

Owttrim GW, Hofmann S, Kuhlemeier C, 1991. Divergent genes for translation initiation factor *eIF-4A* are coordinately expressed in tobacco [J]. *Nucleic Acids Research*, **19** (20): 5491—5496

Owttrim GW, mandel T, Trachsel H *et al.*, 1994. Characterization of the tobacco *eIF-4A* gene family [J]. *Plant Molecular Biology*, **26** (6): 1747—1757

Park W, Li JJ, Song R *et al.*, 2002. *CARPEL FACTORY*, a dicer homolog, and *HEN1*, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana* [J]. *Current Biology*, **12**: 1484—1495

Roack S, Linder P, 2004. DEAD-box proteins: the driving forces behind RNA metabolism [J]. *Nature Reviews Molecular Cell Biology*, **5**: 232—241

Roel GL op den Camp, Kuhlemeier C, 1998. Phosphorylation of tobacco eukaryotic translation initiation factor 4A upon pollen tube germination [J]. *Nucleic Acids Research*, **26** (9): 2058—2062

Seki M, Narusaka M, Abe H *et al.*, 2001. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray [J]. *The Plant Cell*, **13**: 61—72

Tanner NK, Cordin O, Banroques J *et al.*, 2003. The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis [J]. *Molecular Cell*, **11**: 127—138

Vashisht AA, Tuteja N, 2005. Cold stress-induced pea DNA helicase 47 is homologous to *eIF4A* and inhibited by DNA-interacting ligands [J]. *Archives of Biochemistry and Biophysics*, **440** (1): 79—90

Wang YC, Geoffrey D, Purnelle B *et al.*, 2000. Tobacco *VDL* gene encodes a plastid DEAD box RNA helicase and is involved in chloroplast differentiation and plant Morphogenesis [J]. *The Plant Cell*, **12**: 2129—2142

Webster C, Gaut RL, Browning KS *et al.*, 1991. Hypoxia enhances phosphorylation of eukaryotic initiation factor 4A in maize root tips [J]. *Journal of Biological Chemistry*, **266** (34): 23341—23346

Yang Y, Sun ZL, Ge LN *et al.*, 2009. Current progress on DEAD-box RNA helicases in higher plants [D]. Lanzhou: Lanzhou University, Q71

Zhang SS, Tao YS, Zheng YL *et al.*, 2006. Cloning and characterization of DEAD-box RNA helicases gene from Maize (*Zea mays* L.) [J]. *Chinese Journal of Biochemistry and Molecular Biology*, **22** (8): 640—646